JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 58, 373–377 (1977)

Trace Optimization Problems and Generalized Geometric Programming

C. H. SCOTT AND T. R. JEFFERSON

School of Mechanical and Industrial Engineering, University of New South Wales, Kensington, Australia

Submitted by Alex McNabb

The approach of generalized geometric programming is extended to develop a dual version of trace optimization problems. As an application we present a new and simple derivation of the Massieu–Planck extremum principle for quantum statistical equilibrium ensembles.

1. INTRODUCTION

A powerful method of constructing dual problems to convex programming problems is the recently developed theory of generalized geometric programming [3]. In this paper, we exploit and extend this theory to construct a dual problem for optimization problems where the objective and equality constraints are the traces of matrix functions. Apart from its intrinsic mathematical interest, this theory provides a new and simple derivation of the Massieu–Planck extremum principle [1]. This relationship of quantum statistics to generalized geometric programming will allow one to bring to bear the very powerful computational techniques associated with geometric programming. This development will be pursued in a further paper.

In Section 2 we introduce and define a trace transform which is used in Section 3 to construct a dual problem to our primal program. The relevant optimality conditions are derived. In Section 4 we specialize our theory to a particular concave matrix function; the entropy of a quantum statistical ensemble.

2. TRACE TRANSFORM

Define the trace transformation of a function $\operatorname{Tr} f(X)$ with domain C into a matrix function $g(Y)$ with domain D by

$$\begin{aligned} g(Y) &= \sup_{X \in C}\{\operatorname{Tr} XY - \operatorname{Tr} f(X)\} \\ &= \sup_{X \in C}\{\operatorname{Tr}(XY - f(X))\}. \end{aligned} \tag{1}$$

ISSN 0022-247X

For $\{\operatorname{Tr} f: C\}$ convex and closed, the trace transform exists and is both convex and closed.

Equation (1) gives rise to a useful inequality.

$$\operatorname{Tr} XY \leqslant \operatorname{Tr} f(X) + g(Y) \tag{2}$$

with equality for

$$Y \in \partial \operatorname{Tr} f(X) \tag{3}$$

where the equality set is defined by

$$\partial \operatorname{Tr} f(X) = \{Y \mid \operatorname{Tr} f(X) + \operatorname{Tr}((X' - X)\, Y) \leqslant \operatorname{Tr} f(X') \text{ for each } X' \in C\}. \tag{4}$$

This is clearly a convex set.

3. Optimization Problem

Consider the following problem: Minimize

$$\operatorname{Tr} f(X) \qquad \text{over} \qquad X \in C \tag{5}$$

subject to

$$X \in \chi \tag{6}$$

where X is a matrix and χ is a subspace. Define

$$\phi = \inf \operatorname{Tr} f(X) \qquad X \in \chi \cap C. \tag{7}$$

Let $[g(Y): D]$ be the trace transform of $[\operatorname{Tr} f(X): C]$ and

$$S = \{Y \mid \operatorname{Tr}(XY) \geqslant 0 \text{ for all } X \in \chi\}. \tag{8}$$

This is clearly a cone. Then $[g(Y): D]$ are convex and closed. Define

$$\psi = \inf g(Y) \qquad Y \in S \cap D. \tag{9}$$

The problems defined by Eqs. (7) and (9) are of the same type and are termed dual problems. Moreover, this duality is symmetric when $[\operatorname{Tr} f(X): C]$ is convex and closed.

Theorem 1. *If $X \in C$ and $Y \in D$, then*

$$\operatorname{Tr} XY \leqslant g(Y) + \operatorname{Tr} f(X) \tag{10}$$

with equality holding iff

$$Y \in \partial \operatorname{Tr} f(X).$$

Proof. Immediate from the trace inequality (2).

THEOREM 2. *If $X \in \chi \cap C$ and $Y \in S \cap D$, then*

$$g(Y) + \operatorname{Tr} f(X) \geqslant 0 \tag{11}$$

with equality holding iff

$$\operatorname{Tr} XY = 0 \qquad \textit{and} \qquad Y \in \partial \operatorname{Tr} f(X).$$

Proof. Since $X \in \chi$, $Y \in S$, by Theorem 1,

$$g(Y) + \operatorname{Tr} f(X) \geqslant \operatorname{Tr} XY \geqslant 0$$

with equality holding in both these inequalities iff

$$\operatorname{Tr} XY = 0 \qquad \text{and} \qquad Y \in \partial \operatorname{Tr} f(X).$$

Hence the extremality conditions are

$$\begin{aligned} &\text{(i)} && X \in \chi, \\ &\text{(ii)} && \operatorname{Tr}(XY) = 0, \\ &\text{(iii)} && Y \in \partial \operatorname{Tr}(X). \end{aligned} \tag{12}$$

COROLLARY. *If the extremality conditions have a solution X_0 and Y_0 then*

$$\begin{aligned} &\text{(i)} && X_0 \in \chi \cap \partial g(Y_0), \qquad Y_0 \in S \cap \partial \operatorname{Tr} f(X_0), \\ &\text{(ii)} && \phi + \psi = 0. \end{aligned} \tag{13}$$

4. DERIVATION OF THE MASSIEU–PLANCK EXTREMUM PRINCIPLE

In quantum statistics, the equilibrium density matric ρ_0 is the solution of the following problem: Maximize the entropy

$$S = -\operatorname{Tr} \rho \ln \rho \tag{14}$$

subject to the constraints

$$\operatorname{Tr} \rho = 1 \tag{15}$$

and

$$\operatorname{Tr} \rho X_i = a_i \qquad i = 1, \ldots, m; \tag{16}$$

(Boltzmann's constant has been set equal to unity for convenience), where Eq. (16) includes such constraints on the system as energy, number of particles, electron spin measurements, etc. Let

$$C = \{\rho \mid \operatorname{Tr} \rho = 1, \rho \geqslant 0\} \tag{17}$$

and

$$\chi = \{\rho \mid \operatorname{Tr} \rho(X_i - a_i) = 0 \text{ for all } i\}. \tag{18}$$

Hence the entropy maximization problem may be written in the form:

$$\begin{array}{llll} \text{minimize} & \operatorname{Tr} \rho \ln \rho & \text{over} & \rho \in C \\ \text{subject to} & & \rho \in \chi & \end{array} \tag{19}$$

where $\operatorname{Tr} \rho \ln \rho$ is a convex function [2].

We now construct the corresponding dual problem. The trace transform is given by

$$g(\zeta) = \sup_{\rho \in C} \operatorname{Tr}(\rho\zeta - \rho \ln \rho)\}. \tag{20}$$

Taking the first variation of g with respect to ρ results in the supremum being attained when

$$\zeta - 1 - \ln \rho + \lambda = 0$$

where λ is the Lagrange multiplier associated with the constraint $\rho \in C$,

$$\rho_0 = e^{\lambda - 1 + \zeta}.$$

Using constraint (17) to determine λ, we have

$$\rho_0 = \exp \zeta / \operatorname{Tr} \exp \zeta. \tag{21}$$

Hence the trace transform which is our dual objective may be written as

$$g(\zeta) = \ln \operatorname{Tr} \exp \zeta. \tag{22}$$

We now determine ζ from the dual subspace condition, Eq. (8)

$$\operatorname{Tr} \rho\zeta = 0. \tag{23}$$

Equation (23) must be true for all ρ satisfying $\rho \in \chi$. Hence

$$\zeta = \sum_i - \beta_i(X_i - a_i) \tag{24}$$

where β_i are scalar multipliers.

Substituting into Eq. (22) we have that

$$g = \ln \operatorname{Tr} \exp \left(- \sum_i \beta_i(X_i - a_i) \right). \tag{25}$$

The trace inequality, Eq. (2), is given by

$$\ln \operatorname{Tr} \exp \left(- \sum_i \beta_i(X_i - a_i) \right) \geqslant \operatorname{Tr} \left(- \sum_i \beta_i(X_i - a_i)\, \rho - \rho \ln \rho \right) \tag{26}$$

with equality when

$$\begin{aligned}\rho_0 &= \frac{\exp(-\sum_i \beta_i(X_i - a_i))}{\operatorname{Tr}\exp(-\sum_i \beta_i(X_i - a_i))} \\ &= \frac{\exp - \sum_i \beta_i X_i}{\operatorname{Tr}(\exp - \sum_i \beta_i X_i)}.\end{aligned} \tag{27}$$

The extremum principle for Massieu–Planck functions is contained in Eqs. (25), (26), and (27).

References

1. A. Huber, An inequality for traces and its application to extremum principles and related theorems in quantum statistical mechanics, *in* "Methods and Problems of Theoretical Physics" (J. E. Bowcock, Ed.), North-Holland, Holland, 1970.
2. E. Lieb, Some convexity and subadditivity properties of entropy, *Bull. Amer. Math. Soc.* **81** (1975), 1–13.
3. E. L. Peterson, Symmetric duality for generalised geometric programming, *SIAM J. Appl. Math.* **19** (1970), 487–526.